RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/533,466

DATE: 03/29/2001
TIME: 16:16:47

Input Set : A:\21416942.app
Output Set: N:\CRF3\03292001\I533466.raw

<110> APPLICANT: COLLART, FRANK R.
HUBERMAN, ELIEZER
JOACIMIAK, ANDRZEJ
ZHANG, RONGGUANG
WESTBROOK, EDWIN M.
<120> TITLE OF INVENTION: USE OF CRYSTAL STRUCTURE OF BACTERIAL IMP DEHYDROGENASE
TO DESIGN INHIBITORS OF BACTERIAL GROWTH
<130> FILE REFERENCE: 21416/90042
<140> CURRENT APPLICATION NUMBER: 09/533,466
<141> CURRENT FILING DATE: 2000-03-23
<160> NUMBER OF SEQ ID NOS: 23
<170> SOFTWARE: PatentIn Ver. 2.1
<210> SEQ ID NO: 1
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Streptococcus pyogenes
<400> SEQUENCE: 1
Pro Gly Ser Ile Cys Thr Thr Arg Val Val Ala Gly Val Gly Val
1 5 10 15
<210> SEQ ID NO: 2
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Bacillus subtilis
<400> SEQUENCE: 2
Pro Gly Ser Ile Cys Thr Thr Arg Val Val Ala Gly Val Gly Val
1 5 10 15
<210> SEQ ID NO: 3
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli
<400> SEQUENCE: 3
Pro Gly Ser Ile Cys Thr Thr Arg Ile Val Thr Gly Val Gly Val
1 5 10 15
<210> SEQ ID NO: 4
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Bacillus subtilis
<400> SEQUENCE: 4
Pro Gly Ser Ile Cys Thr Thr Arg Val Val Ala Gly Val Gly Val
1 5 10 15
<210> SEQ ID NO: 5
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Mycobacterium tuberculosis
<400> SEQUENCE: 5
Pro Gly Ser Ile Cys Thr Thr Arg Val Val Ala Gly Val Gly Val
1 5 10 15

RAW SEQUENCE LISTING DATE: 03/29/2001
PATENT APPLICATION: US/09/533,466 TIME: 16:16:47

Input Set : A:\21416942.app
Output Set: N:\CRF3\03292001\I533466.raw

<210> SEQ ID NO: 6
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 6
Ser Gly Ser Ile Cys Ile Thr Gln Glu Val Leu Ala Cys Gly Arg
1 5 10 15

<210> SEQ ID NO: 7
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Mus musculus
<400> SEQUENCE: 7
Ser Gly Ser Ile Cys Ile Thr Gln Glu Val Leu Ala Cys Gly Arg
1 5 10 15

<210> SEQ ID NO: 8
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis thaliana
<400> SEQUENCE: 8
Ser Gly Ser Ile Cys Ile Thr Gln Glu Val Leu Ala Cys Gly Arg
1 5 10 15

<210> SEQ ID NO: 9
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Leishmania donovani
<400> SEQUENCE: 9
Ser Gly Ser Ile Cys Ile Thr Gln Glu Val Leu Ala Cys Gly Arg
1 5 10 15

<210> SEQ ID NO: 10
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae
<400> SEQUENCE: 10
Thr Gly Ser Ile Cys Ile Thr Gln Lys Val Met Ala Cys Gly Arg
1 5 10 15

<210> SEQ ID NO: 11
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Drosophila melanogaster
<400> SEQUENCE: 11
Ser Gly Ser Ile Cys Ile Thr Gln Glu Val Met Ala Cys Gly Arg
1 5 10 15

<210> SEQ ID NO: 12
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Streptococcus pyogenes
<400> SEQUENCE: 12
Met Ala Lys Gly Ser Ser Asp Arg Tyr Phe Gln Ser Asp Asn Ala Ala
1 5 10 15

RAW SEQUENCE LISTING DATE: 03/29/2001
PATENT APPLICATION: US/09/533,466 TIME: 16:16:47

Input Set : A:\21416942.app
Output Set: N:\CRF3\03292001\I533466.raw

```
140 Asp Lys Leu Val Pro Glu Gly
141              20
144 <210> SEQ ID NO: 13
145 <211> LENGTH: 23
146 <212> TYPE: PRT
147 <213> ORGANISM: Bacillus subtilis
149 <400> SEQUENCE: 13
150 Met Ser Lys Gly Ser Ser Asp Arg Tyr Phe Gln Ser Asp Asn Ala Ala
151   1               5                  10                  15
153 Asp Lys Leu Val Pro Glu Gly
154              20
157 <210> SEQ ID NO: 14
158 <211> LENGTH: 24
159 <212> TYPE: PRT
160 <213> ORGANISM: Escherichia coli
162 <400> SEQUENCE: 14
163 Met Lys Lys Gly Ser Ser Asp Arg Tyr Phe Gln Gly Ser Val Asn Glu
164   1               5                  10                  15
166 Ala Asn Lys Leu Val Pro Glu Gly
167              20
170 <210> SEQ ID NO: 15
171 <211> LENGTH: 21
172 <212> TYPE: PRT
173 <213> ORGANISM: Bacillus subtilis
175 <400> SEQUENCE: 15
176 Met Glu Lys Gly Ser Lys Asp Arg Tyr Phe Gln Glu Glu Asn Lys Lys
177   1               5                  10                  15
179 Phe Val Pro Glu Gly
180              20
183 <210> SEQ ID NO: 16
184 <211> LENGTH: 30
185 <212> TYPE: PRT
186 <213> ORGANISM: Mycobacterium tuberculosis
188 <400> SEQUENCE: 16
189 Met Arg Gly Arg Gly Gly Ala Thr Ser Tyr Ser Lys Asp Arg Tyr Phe
190   1               5                  10                  15
192 Ala Asp Asp Ala Leu Ser Glu Asp Lys Leu Val Pro Glu Gly
193              20                  25                  30
196 <210> SEQ ID NO: 17
197 <211> LENGTH: 23
198 <212> TYPE: PRT
199 <213> ORGANISM: Homo sapiens
201 <400> SEQUENCE: 17
202 Met Asp Lys His Leu Ser Ser Gln Asn Arg Tyr Phe Ser Glu Ala Asp
203   1               5                  10                  15
205 Lys Ile Lys Val Ala Gln Gly
206              20
209 <210> SEQ ID NO: 18
210 <211> LENGTH: 23
```

RAW SEQUENCE LISTING DATE: 03/29/2001
PATENT APPLICATION: US/09/533,466 TIME: 16:16:47

Input Set : A:\21416942.app
Output Set: N:\CRF3\03292001\I533466.raw

211 <212> TYPE: PRT
212 <213> ORGANISM: Mus musculus
214 <400> SEQUENCE: 18
215 Met Asp Lys His Leu Ser Ser Gln Asn Arg Tyr Phe Ser Glu Ala Asp
216 1 5 10 15
218 Lys Ile Lys Val Ala Gln Gly
219 20
222 <210> SEQ ID NO: 19
223 <211> LENGTH: 27
224 <212> TYPE: PRT
225 <213> ORGANISM: Arabidopsis thaliana
227 <400> SEQUENCE: 19
228 Met Glu Arg Gly Asp Ala Lys Gly Ala Ala Met Ser Arg Tyr Tyr His
229 1 5 10 15
231 Asn Glu Met Asp Lys Met Lys Val Ala Gln Gly
232 20 25
235 <210> SEQ ID NO: 20
236 <211> LENGTH: 27
237 <212> TYPE: PRT
238 <213> ORGANISM: Leishmania donovani
240 <400> SEQUENCE: 20
241 Met Gln Lys Thr Gly Thr Lys Gly Asn Ala Ser Thr Ser Arg Tyr Phe
242 1 5 10 15
244 Ser Glu Ser Asp Ser Val Leu Val Ala Gln Gly
245 20 25
248 <210> SEQ ID NO: 21
249 <211> LENGTH: 21
250 <212> TYPE: PRT
251 <213> ORGANISM: Drosophila melanogaster
253 <400> SEQUENCE: 21
254 Met Thr Lys Gly Ser Asp Gln Arg Tyr Leu Gly Asp Gln Thr Lys Leu
255 1 5 10 15
257 Lys Ile Ala Gln Gly
258 20
261 <210> SEQ ID NO: 22
262 <211> LENGTH: 23
263 <212> TYPE: PRT
264 <213> ORGANISM: Saccharomyces cerevisiae
266 <400> SEQUENCE: 22
267 Met Ser Gln Gly Lys Glu Ser Gly Lys Arg Tyr Leu Ser Glu Asn Glu
268 1 5 10 15
270 Ala Val Gln Val Ala Gln Gly
271 20
274 <210> SEQ ID NO: 23
275 <211> LENGTH: 477
276 <212> TYPE: PRT
277 <213> ORGANISM: Artificial Sequence
279 <220> FEATURE:
280 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic

RAW SEQUENCE LISTING DATE: 03/29/2001
PATENT APPLICATION: US/09/533,466 TIME: 16:16:47

Input Set : A:\21416942.app
Output Set: N:\CRF3\03292001\I533466.raw

```
     281       polypeptide
     283 <220> FEATURE:
     284 <221> NAME/KEY: MOD_RES
     285 <222> LOCATION: (1)..(477)
     286 <223> OTHER INFORMATION: "Xaa" represents selenomethionine
     288 <400> SEQUENCE: 23
     289 Ser Asn Trp Asp Thr Lys Phe Leu Lys Lys Gly Tyr Thr Phe Asp Asp
     290   1               5                  10                  15
     292 Val Leu Leu Ile Pro Ala Glu Ser His Val Leu Pro Asn Glu Val Asp
     293              20                  25                  30
     295 Leu Lys Thr Lys Leu Ala Asp Asn Leu Thr Leu Asn Ile Pro Ile Ile
     296          35                  40                  45
W--> 298 Thr Ala Ala Xaa Asp Thr Val Thr Gly Ser Lys Xaa Ala Ile Ala Ile
     299      50                  55                  60
W--> 301 Ala Arg Ala Gly Gly Leu Gly Val Ile His Lys Asn Xaa Ser Ile Thr
     302  65                  70                  75                  80
     304 Glu Gln Ala Glu Glu Val Arg Lys Val Lys Arg Ser Glu Asn Gly Val
     305                  85                  90                  95
     307 Ile Ile Asp Pro Phe Phe Leu Thr Pro Glu His Lys Val Ser Glu Ala
     308             100                 105                 110
W--> 310 Glu Glu Leu Xaa Gln Arg Tyr Arg Ile Ser Gly Val Pro Ile Val Glu
     311         115                 120                 125
W--> 313 Thr Leu Ala Asn Arg Lys Leu Val Gly Ile Ile Thr Asn Arg Asp Xaa
     314     130                 135                 140
W--> 316 Arg Phe Ile Ser Asp Tyr Asn Ala Pro Ile Ser Glu His Xaa Thr Ser
     317 145                 150                 155                 160
     319 Glu His Leu Val Thr Ala Ala Val Gly Thr Asp Leu Glu Thr Ala Glu
     320                 165                 170                 175
     322 Arg Ile Leu His Glu His Arg Ile Glu Lys Leu Pro Leu Val Asp Asn
     323             180                 185                 190
     325 Ser Gly Arg Leu Ser Gly Leu Ile Thr Ile Lys Asp Ile Glu Lys Val
     326         195                 200                 205
     328 Ile Glu Phe Pro His Ala Ala Lys Asp Glu Phe Gly Arg Leu Leu Val
     329     210                 215                 220
     331 Ala Ala Ala Val Gly Val Thr Ser Asp Thr Phe Glu Arg Ala Glu Ala
     332 225                 230                 235                 240
     334 Leu Phe Glu Ala Gly Ala Asp Ala Ile Val Ile Asp Thr Ala His Gly
     335                 245                 250                 255
     337 His Ser Ala Gly Val Leu Arg Lys Ile Ala Glu Ile Arg Ala His Phe
     338             260                 265                 270
     340 Pro Asn Arg Thr Leu Ile Ala Gly Asn Ile Ala Thr Ala Glu Gly Ala
     341         275                 280                 285
     343 Arg Ala Leu Tyr Asp Ala Gly Val Asp Val Val Lys Val Gly Ile Gly
     344     290                 295                 300
     346 Pro Gly Ser Ile Cys Thr Thr Arg Val Val Ala Gly Val Gly Val Pro
     347 305                 310                 315                 320
     349 Gln Val Thr Ala Ile Tyr Asp Ala Ala Ala Val Ala Arg Glu Tyr Gly
     350                 325                 330                 335
     352 Lys Thr Ile Ile Ala Asp Gly Gly Ile Lys Tyr Ser Gly Asp Ile Val
```

VERIFICATION SUMMARY DATE: 03/29/2001
PATENT APPLICATION: US/09/533,466 TIME: 16:16:48

Input Set : A:\21416942.app
Output Set: N:\CRF3\03292001\I533466.raw

L:298 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:301 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:310 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:313 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:316 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:355 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:361 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:367 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:370 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:373 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23